Handwritten Notes - Mini Project

1. Linear Regression - MSE Cost Function Gradient

Hypothesis :

$$\hat{y}_i = B_0 + B_1 x_i$$

Cost Function (MSE) :

$$J(B_0, B_1) = \frac{1}{m} \sum_{i=1}^{m} (\hat{y}_i - y_i)^2$$

Gradients :

* For B_0 :

$$\frac{\partial J}{\partial B_0} = \frac{2}{m} \sum_{i=1}^{m} (\hat{y}_i - y_i)$$

* For B_1 :

$$\frac{\partial J}{\partial B_1} = \frac{2}{m} \sum_{i=1}^{m} (\hat{y}_i - y_i) x_i$$

Update Rule :

$$B_j = B_j - \alpha \frac{\partial J}{\partial B_j}$$

2. Logistic Regression - Cross - Entropy Loss Gradient

Sigmoid Function :

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Hypothesis :

$$h_\theta(x) = \sigma(\theta^T x)$$

Cost Function (Binary Cross-Entropy):

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[y_i \log(h_\theta(x_i)) + (1-y_i)\log(1-h_\theta(x_i))\right]$$

Gradient:

$$\nabla_\theta J(\theta) = \frac{1}{m}X^T(h_\theta(X) - y)$$

3. Sigmoid Function Derivative

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

Derivative:

$$\frac{d}{dz}\sigma(z) = \sigma(z)(1-\sigma(z))$$

4. Matrix Operations Example

Let: $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $y = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

Compute $X^T y$: $X^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$, $X^T y = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 6 \end{bmatrix}$

$$= \begin{bmatrix} (1)(5) + (3)(6) \\ (2)(5) + (4)(6) \end{bmatrix}$$

$$= \begin{bmatrix} 23 \\ 34 \end{bmatrix}$$

5. Normal Equation Derivation

Cost Function (vectorized):

$$J(\beta) = \frac{1}{m}(X\beta - y)^T (X\beta - y)$$

Setting derivative = 0 :

$$\frac{\partial J}{\partial \beta} = \frac{2}{m} X^T (X\beta - y) = 0$$

Solve :

$$X^T X \beta = X^T y$$

Normal Equation :

$$\beta = (X^T X)^{-1} X^T y$$